

A local lemma via entropy compression

Rogério G. Alves¹
Aldo Procacci²
Remy Sanchis²

¹ Departamento de Matemática UFOP 35400-000 - Ouro Preto - MG Brazil

² Departamento de Matemática UFMG 30161-970 - Belo Horizonte - MG Brazil

February 1, 2017

Abstract

In the framework of the probabilistic method in combinatorics we present a new local lemma based on an entropy compression random algorithm and we show through examples that it can efficiently replace the Moser-Tardos algorithmic version of the Lovász Local Lemma as well as its Pegden extension in a restricted setting where the underlying probability space of the bad events is generated by a finite family of i.i.d. uniform random variables. We then use this new lemma to improve the upper bound for the β -frugal vertex chromatic index of a bounded degree graph recently obtained by Ndreca et al. (Eur J Comb 33: 592-609, 2012).

Keywords: Probabilistic Method in combinatorics; Lovász Local Lemma; Randomized algorithms.

MSC numbers: 05D40, 68W20.

1 Introduction

The Lovász Local Lemma (LLL), originally formulated by Erdős and Lovász [9], is a powerful tool in the framework of the probabilistic method in combinatorics to prove the existence of some combinatorial objects with certain desirable properties (such as a proper coloring of the edges of a graph). The basic idea of the LLL is that the existence of the combinatorial object under analysis is guaranteed once a certain event (the good event) in some probability space has a non zero probability to occur. In turn, the occurrence of this good event is guaranteed if a family of (bad) events do not occur. The LLL then provides a sufficient condition on the probabilities of these bad events to ensure that, with non-zero probability, none of them occur.

1.1 Lovász Local Lemma: the original version

In order to enunciate the Lovász local Lemma we need to introduce the concept of dependency graph of a family of events. Let $G = (V, E)$ be a graph with vertex set V and edge set E . Two vertices $v, v' \in V$ are adjacent if $\{v, v'\} \in E$. Given $v \in V$, let $\Gamma_G^*(v)$ be the set of vertices of G adjacent to v (i.e. the neighborhood of v). We denote $\Gamma_G(v) = \Gamma_G^*(v) \cup \{v\}$.

Definition 1.1 *Given a family of events \mathcal{A} in some probability space, a dependency graph \mathcal{G} for this family is a graph with vertex set \mathcal{A} and edge set \mathcal{E} such that, each event $A \in \mathcal{A}$ is independent of the σ -algebra generated by the collection of events $\mathcal{A} \setminus \Gamma_{\mathcal{G}}(A)$.*

Given a family of events \mathcal{A} in some probability space, let \bar{A} be the complement event of $A \in \mathcal{A}$, so that $\bigcap_{A \in \mathcal{A}} \bar{A}$ is the event that none of the events in the family \mathcal{A} occur.

Hereafter the product over the empty set is equal to one, if n is an integer then $[n]$ denotes the set of integers $\{1, 2, \dots, n\}$ and if X is a set then $|X|$ denotes its cardinality.

Theorem 1.2 (Lovász local Lemma) *Let \mathcal{A} be a finite collection of events in a probability space. Let \mathcal{G} be a dependency graph for the family \mathcal{A} . Let $\boldsymbol{\mu} = \{\mu_A\}_{A \in \mathcal{A}}$ be a collection of real numbers in $[0, +\infty)$. If, for each $A \in \mathcal{A}$,*

$$\text{Prob}(A) \leq \frac{\mu_A}{\Phi_A(\boldsymbol{\mu})} \quad (1.1)$$

with

$$\Phi_A(\boldsymbol{\mu}) = \prod_{A' \in \Gamma_{\mathcal{G}}(A)} (1 + \mu_{A'}) \quad (1.2)$$

then

$$\text{Prob}\left(\bigcap_{A \in \mathcal{A}} \bar{A}\right) > 0$$

i.e. the probability that none of the events in the family \mathcal{A} occur is strictly positive.

The popularity of the Lovász Local Lemma is basically due to the fact that it can be implemented for a wide class of problems in combinatorics, in such a way that the condition (1.1), once some few parameters have been suitably tuned, can be easily checked.

1.2 The improved version of the local lemma via cluster expansion

Recently a surprising connection between the LLL and the cluster expansion of the abstract polymer gas, pointed out by Scott and Sokal [22], has attracted the interest of several researchers in the area of statistical mechanics. More precisely, Scott and Sokal showed that the LLL can be viewed as a reformulation of the Dobrushin criterion [8] for the convergence for the cluster expansion of the hard-core lattice gas on a graph G .

Two years after the work of Scott and Sokal, Fernández and Procacci [11] obtained a sensible improvement of the Dobrushin criterion via cluster expansion methods. Such a result has then been used by Bissacot et al. [4] to obtain an improved version of the LLL. To enunciate this improved version of the LLL we first need the following definition. Given a graph G , a set of vertices Y of G is *independent* if no edge $e = \{v, v'\}$ of G is such that $v \in Y$ and $v' \in Y$.

Theorem 1.3 (Cluster Expansion Local Lemma (CELL)) *In the same hypothesis of Theorem 1.2, if, for each $A \in \mathcal{A}$,*

$$\text{Prob}(A) \leq \frac{\mu_A}{\Psi_A(\boldsymbol{\mu})} \quad (1.3)$$

with

$$\Psi_A(\boldsymbol{\mu}) = \sum_{\substack{Y \subseteq \Gamma_{\mathcal{G}}(A) \\ Y \text{ independent}}} \prod_{A' \in Y} \mu_{A'} \quad (1.4)$$

then

$$\text{Prob}\left(\bigcap_{A \in \mathcal{A}} \bar{A}\right) > 0$$

The improvement on Theorem 1.2 is immediately recognized by noting that

$$\Phi_A(\boldsymbol{\mu}) = \prod_{A' \in \Gamma_{\mathcal{G}}(A)} (1 + \mu_{A'}) = \sum_{Y \subseteq \Gamma_{\mathcal{G}}(A)} \prod_{A' \in Y} \mu_{A'} \geq \sum_{\substack{Y \subseteq \Gamma_{\mathcal{G}}(A) \\ Y \text{ indep in } \mathcal{G}}} \prod_{A' \in Y} \mu_{A'} = \Psi_A(\boldsymbol{\mu})$$

Remark. In general it might be not easy to compute the function $\Psi_A(\boldsymbol{\mu})$ defined in (1.4). However, when $\Gamma_{\mathcal{G}}(A)$ is the union of m cliques $c_1(A), \dots, c_m(A)$, then setting

$$\Xi_A(\boldsymbol{\mu}) = \prod_{i=1}^k \left[1 + \sum_{A' \in c_i(A)} \mu_{A'} \right] \quad (1.5)$$

we have

$$\Psi_A(\boldsymbol{\mu}) = \sum_{\substack{Y \subseteq \Gamma_{\mathcal{G}}(A) \\ Y \text{ independent}}} \prod_{A' \in Y} \mu_{A'} \leq \prod_{i=1}^k \left[1 + \sum_{A' \in c_i} \mu_{A'} \right] = \Xi_A(\boldsymbol{\mu}) \quad (1.6)$$

The above remark allows us to enunciate the so called “clique approximation” of the CELL which will be explicitly used Section 2.1 ahead.

Theorem 1.4 (Clique approximation of CELL) *In the same hypothesis of Theorem 1.2 suppose that the neighborhood $\Gamma_{\mathcal{G}}(A)$ in \mathcal{G} is the union of $c_1(A), \dots, c_m(A)$ cliques. If, for each $A \in \mathcal{A}$,*

$$Prob(A) \leq \frac{\mu_A}{\Xi_A(\boldsymbol{\mu})} \quad (1.7)$$

where $\Xi_A(\boldsymbol{\mu})$ is defined in (1.5), then

$$Prob\left(\bigcap_{A \in \mathcal{A}} \bar{A}\right) > 0$$

Of course, Theorem 1.4 is weaker than Theorem 1.3 but generally¹ is stronger than Theorem 1.2 and much easier to implement than Theorem 1.3. So much so that, as far as we know, Theorem 1.4 is until today the only tool² that has been used to improve on previous bounds via LLL. Namely, via Theorem 1.4, better estimates for latin transversal (see [4]) and for several chromatic indices of bounded degree graphs (see [18] and [5]) have been obtained.

1.3 The algorithmic version of the local lemma

Despite its popularity, the LLL has been object of a recurring criticism about its inherently non-constructive character. Specifically, the LLL provides sufficient conditions for the probability that none of the undesirable events in some suitable probability space occur to be strictly positive and this implies the existence of at least one outcome in the probability space which realizes the occurrence of the “good” event. However the LLL, in its original version of Theorem 1.2, as well as in the improved version of Theorem 1.3, does not provide any algorithm able to find such a configuration. The efforts to devise an algorithmic version of the LLL, going back to the work of Beck [3], Alon [1] and others (see e.g. reference list in [16]), finally culminated with the breakthrough work by Moser and Tardos [15, 17], who presented a fully algorithmic version

¹Theorem 1.4 can be weaker than the original LLL Theorem 1.2 when the identified cliques $c_1(A), \dots, c_m(A)$ are too much overlapped. This happens for example in the case of perfect and separating hash families (see [20]).

²With the exception of [20] where the full CELL has been used to improve bounds for Hash families.

of LLL in a specific “variable setting” which however covers basically all known applications of LLL.

The variable setting. Let \mathcal{V} be a finite family of mutually independent random variables and let Ω be the probability space determined by these variables. As usual, if $X \subset \mathcal{V}$, then $\Omega|_X$ will denote the restriction of the sample space Ω to the variables in X and if $\omega \in \Omega$, we denote by $\omega|_X$ the restriction (or projection) of the configuration ω to the restricted space $\Omega|_X$.

Let \mathcal{A} be a finite family of events on Ω , so that each $A \in \mathcal{A}$ has a given probability $Prob(A)$ to occur and depends in principle on all variables in \mathcal{V} . Moser and Tardos made the critical assumption that each $A \in \mathcal{A}$ depends actually only on some subset of the random variables of the family \mathcal{V} and they denoted by $vbl(A)$ the minimal (with respect to inclusion) subset of \mathcal{V} on which A depends. In other words, to decide whether or not the event A occurs, it is sufficient to look only at the value of the variables in $vbl(A)$. In particular, the event A is independent of the σ -algebra generated by the variables in $\mathcal{V} \setminus vbl(A)$. Since variables in \mathcal{V} are assumed to be mutually independent, any two events $\{A, A'\} \in \mathcal{A}$ such that $vbl(A) \cap vbl(A') = \emptyset$ are necessarily independent. Therefore the family \mathcal{A} has a natural dependency graph, i.e. the graph \mathcal{G} with vertex-set \mathcal{A} and edge-set \mathcal{E} constituted by the pairs $\{A, A'\} \subset \mathcal{A}$ such that $vbl(A) \cap vbl(A') \neq \emptyset$. In this framework Moser and Tardos defined the following algorithm.

MT-Algorithm.

- Step 0. Choose a random evaluation of all the variables in \mathcal{V} .
- Step i . If some of the events of the family \mathcal{A} occur, select one of them (at random or according to some deterministic rule), say A , and take a new evaluation (resampling) only of the variables in $vbl(A)$, keeping unchanged all the other variables in $\mathcal{V} \setminus vbl(A)$.

The algorithm stops when an evaluation of the variables in \mathcal{V} is reached such that none of the events in the family \mathcal{A} occur.

Within this “variable setting” Moser and Tardos proved [17] an algorithmic version of the LLL identical to Theorem 1.2, They also proved that their MT-algorithm finds an evaluation of the variable in \mathcal{V} such that none of the bad events of the family \mathcal{A} occur in an expected time proportional to the number of events in the family \mathcal{A} . This algorithmic version of the LLL proposed by Moser and Tardos has been very recently improved by Pegden [19] by replacing condition (1.1) with condition (1.8) using once again the connection with cluster expansion (see also [2] and [13]). Pegden’s result is as follows.

Theorem 1.5 (Algorithmic CELL) *Let \mathcal{V} be a finite set of mutually independent random variables and let Ω be the probability space generated by these variables. Let \mathcal{A} be a finite set of events in Ω with each $A \in \mathcal{A}$ depending on a subset $vbl(A)$ of the variables \mathcal{V} . Let \mathcal{G} be the natural dependency graph for the family \mathcal{A} . Let $\boldsymbol{\mu} = \{\mu_A\}_{A \in \mathcal{A}}$ be a collection of real numbers in $[0, +\infty)$. If, for each $A \in \mathcal{A}$,*

$$Prob(A) \leq \frac{\mu_A}{\Psi_A(\boldsymbol{\mu})} \tag{1.8}$$

with

$$\Psi_A(\boldsymbol{\mu}) = \sum_{\substack{Y \subseteq \Gamma_{\mathcal{G}}(A) \\ Y \text{ independent}}} \prod_{A' \in Y} \mu_{A'} \tag{1.9}$$

then there is an evaluation of the variables \mathcal{V} such that none of the events in the family \mathcal{A} occur and the MT-algorithm finds such an evaluation in an expected total number of steps equal to $\sum_{A \in \mathcal{A}} \mu_A$.

1.4 The objective of the present paper

Theorem 1.5 is a significant advance, but it is not the end of the story. Due to the simplicity of the MT-algorithm, the question has been raised as to whether it is possible to design some more sophisticated and possibly more efficient algorithm, eventually depending on the specific problem treated, able to improve the bounds given by the LLL or CELL. These ideas have been very recently developed by several authors (see e.g. [6, 10, 21] and references therein) and new algorithms have been implemented for some specific graph coloring problems to obtain bounds which are sensibly better than those given by LLL and/or CELL. In particular, Esperet and Parreau [10] devised an algorithm able to obtain new bounds for the chromatic indices of the acyclic edge coloring and star coloring of a graph with maximum degree Δ . Moreover the same authors suggested that their algorithm based on *entropy compression*³ is sufficiently general to treat most of the applications in graph coloring problems covered by the LLL.

In this note we show that if one further restricts the Moser-Tardos variable setting by imposing that the random variables of the set \mathcal{V} are not only independent but also uniformly identically distributed and taking values in a finite set, then, by developing the ideas illustrated in [10], it is possible to formulate a criterion, that we have called “Entropy Compression Lemma” which is alternative to either the LLL or CELL, and whose implementation for a huge class of problems in combinatorics (including problems not related with graph coloring) is, we think, even more straightforward than that of the LLL and CELL. We also show through known examples that this new Lemma generally improves on the LLL and the CELL. We finally apply the Entropy Compression Lemma to obtain an upper bound for the β -frugal chromatic index of a graph with maximum degree Δ which improves the best known bound given recently in [18].

The rest of the paper is organized as follows. In Section 2 we introduce the notations and we present our main result (Theorem 2.3). In Section 3 we give the proof of Theorem 2.3. Finally in Section 4 we present an application about coloring graphs frugally.

2 Notations and results: the “Entropy Compression Lemma”

The i.i.d. uniform variable setting. As in the Moser-Tardos framework, let $\mathcal{V} = \{x_1, \dots, x_N\}$ be a set of $N \in \mathbb{N}$ mutually independent random variables on a common probability space. We further require that each random variable x_i in \mathcal{V} takes values in the same set of integers $[k]$ according to the uniform distribution. In other words the variables $\{x_1, \dots, x_N\}$ are independent and identically distributed (i.i.d.). With these assumptions the sample space generated by the variables in \mathcal{V} is $\Omega = [k]^N$ so that an outcome $\omega \in \Omega$ is just an ordered N -tuple $\omega = (k_1, \dots, k_N)$ with $k_i \in [k]$ for all $i \in [N]$. The probability space determined by the variables \mathcal{V} is the triple $(\Omega, \mathcal{F}, \mu)$ where $\mathcal{F} = 2^\Omega$ is the σ -algebra and, for any event $A \in \mathcal{F}$, $\mu(A) = |A|/k^N$ is its probability ($|A|$ is the number of outcomes which realize A).

We define the set $E_{\mathcal{A}} \subset \mathbb{N}$ as

$$E_{\mathcal{A}} = \{l \in \mathbb{N} : \exists A \in \mathcal{A} \text{ s.t. } |vbl(A)| = l\} \quad (2.1)$$

and, for $l \in E_{\mathcal{A}}$ we let $\mathcal{A}_l = \{A \in \mathcal{A} : |vbl(A)| = l\}$. I.e. \mathcal{A}_l is the subfamily of \mathcal{A} containing all the events of size l so that the family \mathcal{A} is the disjoint union of sub-families $\{\mathcal{A}_l\}_{l \in E_{\mathcal{A}}}$.

For $y \in \mathcal{V}$, we further denote by $\mathcal{A}(y)$ the subfamily of \mathcal{A} given by $\mathcal{A}(y) = \{A \in \mathcal{A} : y \in vbl(A)\}$ and shortly, for $y \in \mathcal{V}$ and $l \in E_{\mathcal{A}}$, $\mathcal{A}_l(y) = \mathcal{A}_l \cap \mathcal{A}(y)$.

³The expression “entropy compression” in reference to the Moser Tardos algorithm and its variants was probably first used by Terence Tao in a note about Moser-Tardos method he published in his Blog in 2009 [23].

Definition 2.1 Let $A \in \mathcal{A}$ be an event depending on variables $vbl(A) \subset \mathcal{V}$ and let $X \subsetneq vbl(A)$. We say that X is a “seed” of A if, for any two realizations $\omega, \tilde{\omega} \in \Omega|_{vbl(A)}$ of A , $\omega|_X = \tilde{\omega}|_X$ implies $\omega = \tilde{\omega}$ and no $Y \subsetneq X$ has this property.

In other words a proper subset $X \subset vbl(A)$ is called a seed for A if, for any realization $\omega \in \Omega|_{vbl(A)}$ of A , the knowledge of $\omega|_X$ it is sufficient to reconstruct uniquely the full realization ω and no $Y \subsetneq X$ has this property. Note also that if A is elementary, then $X = \emptyset$ is the unique seed of A .

Definition 2.2 An event $A \in \mathcal{A}$ is said recordable if for all $y \in vbl(A)$ there exists $X \subset vbl(A) \setminus \{y\}$ which is a seed of A .

Remark. Note any elementary event A is always recordable since one can choose for any $y \in vbl(A)$, $X = \emptyset$ as seed of A contained in $vbl(A) \setminus \{y\}$. Therefore any event which is not recordable can always be seen as the union of smaller (in the sense of inclusion) recordable events (ultimately as union of elementary events). We will thus suppose hereafter, without loss of generality, that all events in the family \mathcal{A} are recordable.

We need now to introduce now some further definitions which will allow us to describe a general algorithm, called EC-algorithm (from Entropy-Compression), able to find an evaluation of the variables \mathcal{V} such that no bad event in the family \mathcal{A} occurs.

Given $y \in \mathcal{V}$ and $A \in \mathcal{A}(y)$, we define

$$\kappa_y(A) = \min\{|X| : X \subset vbl(A) \setminus \{y\} \text{ and } X \text{ is a seed of } A\} \quad (2.2)$$

Note that, for any $y \in \mathcal{V}$ and $A \in \mathcal{A}_l(y)$, $\kappa_y(A) \in \{0, 1, \dots, l-1\}$ and the case $\kappa_y(A) = 0$ only happens when the event A is elementary.

Let $y \in \mathcal{V}$. We define $K_l(y)$ and K_l as the following sets of integers.

$$K_l(y) = \left\{ \kappa \in \{0, 1, \dots, l-1\} : \kappa_y(A) = \kappa \text{ for some } A \in \mathcal{A}_l(y) \right\}$$

$$K_l = \bigcup_{y \in \mathcal{V}} K_l(y) \quad (2.3)$$

Note that $K_l \subset \{0, 1, \dots, l-1\}$. Once we have the set of integers $E_{\mathcal{A}}$ and, for each $l \in E_{\mathcal{A}}$, the set of integers K_l we define, for $x \in \mathbb{R}$,

$$\phi_{E_{\mathcal{A}}}(x) = 1 + \sum_{l \in E_{\mathcal{A}}} \sum_{\kappa \in K_l} x^{l-\kappa} \quad (2.4)$$

We finally define, for $\kappa \in K_l$, the set

$$\mathcal{A}_l^\kappa(y) = \{A \in \mathcal{A}_l(y) : \kappa_y(A) = \kappa\}$$

and the integer

$$d_l^\kappa = \max_{y \in \mathcal{V}} |\mathcal{A}_l^\kappa(y)| \quad (2.5)$$

Therefore, the number of events $A \in \mathcal{A}$ of size l sharing a fixed variable y in \mathcal{V} and such that $\kappa_y(A) = \kappa$ is bounded above by d_l^κ uniformly in $y \in \mathcal{V}$.

We are now ready to describe the EC-algorithm, which, as we will see in the next section, is the key ingredient in order to prove the main result of this note, i.e. Theorem 2.3 below. Hereafter

we will assume that a total order is fixed on the set of variables \mathcal{V} as well as on the set of events \mathcal{A} . We also choose, for each $y \in \mathcal{V}$ and $A \in \mathcal{A}(y)$, a unique subset $G(A, y) \subset vbl(A) \setminus \{y\}$ such that $G(A, y)$ is a seed of A and $|G(A, y)| = \kappa_y(A)$ (if $\kappa_y(A) = 0$, then $G(A, y) = \emptyset$). We will often say that a variable $x \in \mathcal{V}$ is *colored* if a value in $[k]$ has been assigned to x . In other words, to color a variable $x \in \mathcal{V}$ means to select for it a value $s \in [k]$. It is then clear that to uncolor a colored variable x means that the assigned value in $[k]$ of the variable x is withdrawn and the variable is left undetermined.

EC-Algorithm.

- Step i . Let y be the smallest variable (in the total order chosen) uncolored (i.e. with no assigned value) and take a random evaluation of the value of this variable y (i.e. color this variable).

- i_1) If after the coloring of the variable y no bad event $A \in \mathcal{A}$ occurs, then go to the step $i + 1$.
- i_2) If on the contrary after the coloring of the value of the variable y some set of bad events, say $\mathcal{S}_i \subset \mathcal{A}$, occurs, then necessarily $\mathcal{S}_i \subset \mathcal{A}(y)$. According to the total order previously fixed in the family \mathcal{A} , select the event A which is the smallest in the set \mathcal{S}_i . Such event A will belong to the set $\mathcal{A}_l^\kappa(y)$ for some $l \in E_{\mathcal{A}}$ and some $\kappa \in K_l$. Now uncolor all variables in $vbl(A)$ except the κ variables belonging to the set $G(A, y)$ previously introduced (in other words, uncolor all the $l - \kappa$ variables in $vbl(A) \setminus G(A, y)$). Then go to the step $i + 1$.

- The algorithm stops when all variables has been evaluated (i.e. colored) and no event of the family \mathcal{A} occurs.

We are now in the position to state the main theorem of this note.

Theorem 2.3 (Entropy-compression Lemma) *Let \mathcal{V} be a finite set of i.i.d. uniform random variables taking value in $[k]$. Let \mathcal{A} be a finite family of events depending on these variables. If*

$$k > \left[\inf_{x>0} \frac{\phi_{E_{\mathcal{A}}}(x)}{x} \right] \max_{\substack{l \in E_{\mathcal{A}} \\ \kappa \in K_l}} \left\{ (d_l^\kappa)^{\frac{1}{1-\kappa}} \right\} \quad (2.6)$$

then there is an evaluation of the variables \mathcal{V} such that none of the events in the family \mathcal{A} occur and the EC-algorithm finds such an evaluation, almost surely in an expected running time linear in the number of variables.

Remark 1. Note that

$$\inf_{x>0} \frac{\phi_{E_{\mathcal{A}}}(x)}{x} = \phi'_{E_{\mathcal{A}}}(\tau) \quad (2.7)$$

where τ the unique positive root of the equation $\phi_{E_{\mathcal{A}}}(x) - x\phi'_{E_{\mathcal{A}}}(x) = 0$.

Remark 2. The EC-algorithm still finds an evaluation of the variables \mathcal{V} such that none of the events in the family \mathcal{A} occur even if we only demand k to be greater or equal (instead of strictly greater) than the l.h.s. of (2.6), but in this case we have no control on the expected running time.

2.1 Comparison and examples

It is instructive and illuminating, we think, to start by comparing Theorem 2.3 with the clique approximation of the algorithmic Cluster expansion Local Lemma, namely, Theorem 1.4. In fact, as we will see below, in the i.i.d. uniform variable setting the function $\Xi_A(\boldsymbol{\mu})$ defined in (1.5) can be computed explicitly and it turns to exhibit a structure such that the comparison between conditions (1.7) and (2.6) becomes manifest pointing out very clearly how and when Theorem 2.3 beats Theorem 1.4, which, we repeat, has been the tool used in recent times to improve bounds previously obtained via the classical version of LLL given by Theorem 1.2.

Entropy Compression Lemma versus Clique approximation of the CELL

For the benefit of the reader we start by recalling that in the i.i.d. uniform variable setting it is given a finite set \mathcal{V} of N i.i.d. uniform random variables taking value in $[k]$ and a family of events \mathcal{A} in the probability space determined by the variables \mathcal{V} . Each event $A \in \mathcal{A}$ depends actually on a subset $vbl(A) \subset \mathcal{V}$ and the dependency graph for the family \mathcal{A} is the graph \mathcal{G} with vertex set \mathcal{A} and edge set \mathcal{E} formed by those $\{A, A'\}$ such that $vbl(A) \cap vbl(A') \neq \emptyset$. Let now $E_{\mathcal{A}}$ and K_l be the sets defined in (2.1) and (2.3) respectively.

Let $l \in E_{\mathcal{A}}$. For any event $A \in \mathcal{A}_l$ let us define, recalling definition (2.2), the integer $\kappa_A \in K_l$ as follows

$$\kappa_A = \min_{y \in vbl(A)} \kappa_y(A)$$

Given $(l, \kappa) \in E_{\mathcal{A}} \times K_l$, let us refer to the events $A \in \mathcal{A}$ such that $A \in \mathcal{A}_l$ and $\kappa_A = \kappa$ as an “event of size l and type κ ”. If A is an event of size l and type κ , then, by definition, there is a seed X of A such that $|X| = \kappa$ and, recalling Definition 2.1, this immediately implies that

$$Prob(A) \leq \frac{1}{k^{l-\kappa}}$$

Moreover, recalling definition (2.5), the integer d_l^κ is, for any $(l, \kappa) \in E_{\mathcal{A}} \times K_l$, an upper bound for the number of events of size l and of type κ containing a fixed variable. This implies that the neighborhood $\Gamma_{\mathcal{G}}(A)$ of an event A of size l and type κ is the union of l cliques (one clique for each variables in $vbl(A)$) and, for any $m \in E_{\mathcal{A}}$ and any $\lambda \in K_m$, each of these cliques contains at most d_m^λ events of size m and type λ .

In order to set the condition (1.7), let $\{\mu_{l,\kappa}\}_{(l,\kappa) \in E_{\mathcal{A}} \times K_l}$ be a set of positive numbers and let us chose the set $\{\mu_A\}_{A \in \mathcal{A}}$ in condition (1.7) as follows

$$\mu_A = \mu_{l,\kappa} \quad \text{whenever} \quad A \in \mathcal{A}_l \quad \text{and} \quad \kappa_A = \kappa$$

In other words, we associate to all events of the same size l and of the same type κ the same parameter $\mu_{l,\kappa}$. With these notations and definitions the function $\Xi_A(\boldsymbol{\mu})$ defined in (1.5) is such that

$$\Xi_A(\boldsymbol{\mu}) \leq \left[1 + \sum_{m \in E_{\mathcal{A}}} \sum_{\lambda \in K_m} d_s^\lambda \mu_{m,\lambda} \right]^l. \quad (2.8)$$

Let us now set, for all $(l, \kappa) \in E_{\mathcal{A}} \times K_l$,

$$\mu_{l,\kappa} = x^{l-\kappa} / d_l^\kappa$$

with $x > 0$ being a parameter to be optimized later. Then

$$\Xi_A(\boldsymbol{\mu}) \leq \left[1 + \sum_{m \in E_{\mathcal{A}}} \sum_{\lambda \in K_m} d_s^\lambda \mu_{m,\lambda} \right]^l = \left[1 + \sum_{m \in E_{\mathcal{A}}} \sum_{\lambda \in K_m} x^{m-\lambda} \right]^l = [\phi_{E_{\mathcal{A}}}(x)]^l$$

where, surprising as it may seem, $\phi_{E_A}(x)$ is the very same function defined in the statement of Theorem 2.3.

Therefore condition (1.8) of Theorem 1.5 is satisfied if the following inequality holds for all $(l, \kappa) \in E_A \times K_l$.

$$\frac{1}{k^{l-\kappa}} \leq \frac{1}{d_l^\kappa} \frac{x^{l-\kappa}}{[\phi_{E_A}(x)]^l} \quad \text{for all } (l, \kappa) \in E_A \times K_l \quad (2.9)$$

The condition (2.9) above can be rewritten as follows.

$$k \geq \max_{l, \kappa} \left[\inf_{x>0} \frac{[\phi_E(x)]^{\frac{l}{l-\kappa}}}{x} \right] (d_l^\kappa)^{\frac{1}{l-\kappa}} \quad (2.10)$$

Note that inequality (2.10) is notably similar to the inequality (2.6) which, once satisfied, guarantees that the thesis of Theorem 2.3 is true.

Since for all $(l, \kappa, x) \in E_A \times K_l \times (0, \infty)$ we have that $\frac{l}{l-\kappa} > 1$ and $\phi_E(x) > 1$, it holds

$$\frac{[\phi_{E_A}(x)]^{\frac{l}{l-\kappa}}}{x} \geq \frac{\phi_{E_A}(x)}{x} \quad \text{for all } (l, \kappa) \in E_A \times K_l \text{ and for all } x > 0$$

and hence, for all $(l, \kappa) \in E_A \times K_l$, we have

$$\inf_{x>0} \frac{[\phi_{E_A}(x)]^{\frac{l}{l-\kappa}}}{x} \geq \inf_{x>0} \frac{[\phi_{E_A}(x)]}{x} \quad (2.11)$$

and the equality in (2.11) only holds if $\kappa = 0$.

Therefore Condition (2.6) is always better than Condition (2.10) and these two Conditions coincide only when $K_l = \{0\}$ for all l . In other words Theorem 2.3 is never beaten by Theorem 1.4 and Theorem 2.3 beats Theorem 1.4 every time the maximum of $(d_l^\kappa)^{1/(l-\kappa)}$ is attained at a value (l, κ) with $\kappa \neq 0$.

Let us conclude this section by illustrating how straightforwardly and how generally Theorem 2.3 can be applied by considering three well known and, hopefully, pedagogical examples two of which being not graph coloring problems.

Property B for m -uniform and n -regular Hypergraphs

Consider a finite set V and let $P_2(V) = \{U \subset V : |U| \geq 2\}$. Then a (loopless) hypergraph is a pair $H = (V, E)$ where $E \subset P_2(V)$; the elements of V are called vertices of the hypergraph and the elements of E are called edges of the hypergraph. A hypergraph $H = (V, E)$ is m -uniform if every edge $f \in E$ contains exactly m vertices and it is n -regular if every vertex $v \in V$ is contained in exactly n edges. A hypergraph $H = (V, E)$ has the property B if there is a coloring of V by two colors such that no edge $f \in E$ is monochromatic and, if $H = (V, E)$ is m -uniform and n -regular and has the property B, then we say that the pair (m, n) is good.

Using LLL (i.e Theorem 1.2) Erdős and Lovász [9] showed that a m -uniform and n -regular hypergraph has the property B (i.e. the pair (m, n) is good) if

$$2^{m-1} \geq e[m(n-1) + 1] \quad (2.12)$$

which implies that the pair $(9, 11)$ is good⁴.

⁴Actually the best result is due to McDiarmid [14] who proved that the pair $(8, 12)$ is good by using in a very clever way a ‘lopsided’ variant of the Lovász Local Lemma

Let us now show how to use Theorem 2.3 to improve the result given in (2.12).

Suppose that we color the vertices of a m -uniform and n -regular hypergraph $H = (V, E)$ by choosing at random independently and uniformly from a set of two colors.

We have then a set $\mathcal{V} = \{x_v\}_{v \in V}$ of i.i.d. uniform random variables and each variable x_v takes two values corresponding to the two possible colors (i.e. $k = 2$).

Given an edge $f = \{v_1, \dots, v_m\} \in E$, let A_f be the event that all vertices in f receive the same color. In other words $A_f = \{x_{v_1} = x_{v_2} = \dots = x_{v_m}\}$. We thus have a family of (bad) events $\mathcal{A} = \{A_f\}_{f \in E}$.

Theorem 2.3 tells us that when condition (2.6) is satisfied (with $k = 2$), then there is a two-coloring of the vertex set V of H such that none of the events in the family \mathcal{A} occur, namely, there is a coloring such that no edge of H is monochromatic, i.e., H has the property B.

Now observe that for any $A_f \in \mathcal{A}$ we have that $vbl(A_f) = \{x_v\}_{v \in f}$ and thus, recalling that H is m -uniform, $|vbl(A_f)| = m$. Therefore we simply have $E_{\mathcal{A}} = \{m\}$. Moreover for any $v \in V$ and for any $A \in \mathcal{A}_m(x_v)$ we have that $\kappa_{x_v}(A) = 1$ and hence $K_m = \{1\}$.

Therefore the function $\phi_{E_{\mathcal{A}}}(x)$ defined in (2.4) is in the present case $\phi_{E_{\mathcal{A}}}(x) = 1 + x^{m-1}$ so that

$$\inf_{x>0} \left[\frac{\phi_{E_{\mathcal{A}}}(x)}{x} \right] = \frac{(m-1)}{(m-2)^{\frac{m-2}{m-1}}} \quad (2.13)$$

Finally, recalling that H is n -regular, the number d_m^1 which represents the maximum number of events of size m containing a fixed vertex (i.e the number of edges of H containing a fixed vertex), is simply

$$d_m^1 = n \quad (2.14)$$

We have now all the ingredients to check condition (2.6). Namely, there is a two-coloring of the vertex set V of H such that no edge is monochromatic as soon as

$$2 \geq \frac{(m-1)}{(m-2)^{\frac{m-2}{m-1}}} n^{\frac{1}{m-1}} \quad (2.15)$$

i.e., as soon as

$$2^{m-1} \geq \left[1 + \frac{1}{m-2} \right]^{m-2} (m-1)n \quad (2.16)$$

Condition (2.16) tells us that the pair (9, 12) is good.

Finally let us observe that using judiciously Theorem 1.4 one obtains the condition

$$2^{m-1} \geq \left[\frac{1}{(m-1)n} + \frac{m(n-1)}{n(m-1)} \left(1 + \frac{1}{m-1} \right)^{m-1} \right] (m-1)n$$

which is clearly worst than (2.16) but better than (2.12). In fact, it is sufficiently better than (2.12) to also allow to conclude that the pair (9, 12) is good.

Independent sets.

Let G be a graph with vertex set V , edge set E and with maximum degree Δ . Consider a partition of the vertex set V of G as follows: $V = \cup_{i=1}^n V_i$ with $|V_i| \geq k$. We want to find the least k such that there exists an independent subset I of V with exactly one vertex in each V_i . Without loss of generality, we can assume that G is such that $|V_i| = k$ for all sets V_i . The general case $|V_i| \geq k$ follows using the graph induced by G on a union of n subsets of cardinality k , each of them a subset of one V_i .

Let us explain the i.i.d. variable setting for the present problem. We start by choosing, for each $i \in [n]$, a bijection $f_i : [k] \rightarrow V_i$. Then consider the set of i.i.d random variables $\mathcal{V} = \{x_i\}_{i \in [n]}$ taking value in $[k]$, in such a way that when x_i takes the values $j \in [k]$, then the vertex $v = f_i(j)$ is selected in the set V_i . Let now for each $\{i, i'\} \subset [n]$, $A_{\{i, i'\}}$ be the following event in the probability space generated by variables \mathcal{V} :

$$A_{\{i, i'\}} = \left\{ x_i = j, x_{i'} = j' \quad \text{and} \quad \{f_i(j), f_{i'}(j')\} \in E \right\}$$

In other words $A_{\{i, i'\}}$ is the event that we select the vertex v in V_i and the vertex v' in $V_{i'}$ and $\{v, v'\}$ is an edge of G .

Consider now the family of events $\mathcal{A} = \{A_{\{i, i'\}}\}_{\{i, i'\} \subset [n]}$. Clearly any outcome $\omega \in [k]^n$ such that none of the events in the family \mathcal{A} occur gives an independent set of G .

In the present case we have $vbl(A_{\{i, i'\}}) = \{x_i, x_{i'}\}$ so any event in \mathcal{A} has size $l = 2$. Moreover since $A_{\{i, i'\}}$ are elementary events, we have $\kappa_y(A) = 0$ for any $y \in \mathcal{V}$ and any $A \in \mathcal{A}$. This means that $E_{\mathcal{A}} = \{2\}$ and $K_2 = \{0\}$, so that

$$\inf_{x>0} \frac{\phi_{E_{\mathcal{A}}}(x)}{x} = 2$$

Finally d_2^0 represents in the present case the number of events $A_{\{i, i'\}}$ containing a fixed variable. Since a variable selects k vertices of G and each of these vertices is adjacent to at most Δ other vertices of G , we get that $d_2^0 \leq k\Delta$. Therefore condition (2.6) is written as follows

$$k \geq 2(k\Delta)^{\frac{1}{2}}$$

which is to say $k \geq 4\Delta$. This bound is better than that obtained with the LLL and equal to that obtained with the CELL (see e.g [4]). In this case Theorem 2.3 gives the same bounds as CELL because $K_l = \{0\}$ for any $l \in E_{\mathcal{A}}$. In the next example, where $K_l \neq \{0\}$ for some $l \in E_{\mathcal{A}}$, Theorem 2.3 beats CELL.

Van der Waerden numbers.

Given two integers m and c , the van der Waerden number $W(c, m)$ is the least integer N such that if the set $\{1, 2, \dots, N\}$ is colored using c colors then there is a monochromatic arithmetic progression with m terms.

We have N random variables i.i.d. $\mathcal{V} = \{x_1, \dots, x_N\}$ each varying in the set $[c]$. Let \mathcal{P} denotes the set of all arithmetic progressions with m terms in $[N]$ and let, for $p \in \mathcal{P}$, A_p be the event that the m -term progression p is monochromatic. Consider now the family of bad events $\mathcal{A} = \{A_p\}_{p \in \mathcal{P}}$. So, all bad events are of size m , i.e. $l = m$ and clearly we have $\kappa_y(A) = 1$ for any $y \in \mathcal{V}$ and any $A \in \mathcal{A}$ and thus $E_{\mathcal{A}} = \{m\}$, $K_m = 1$ and therefore

$$\inf_{x>0} \frac{\phi_{E_{\mathcal{A}}}(x)}{x} = \frac{m-1}{(m-2)^{\frac{m-2}{m-1}}} = \left(1 + \frac{1}{m-2}\right)(m-2)^{\frac{1}{m-1}}$$

Moreover we have $d_m^1 \leq m \cdot \lfloor \frac{N}{m} \rfloor \leq N$ (this is the number of arithmetic progressions with m terms in $[N]$ containing a fixed number). We have now all the ingredients to apply Theorem 2.3. So we get that, whatever the coloring is, there is no monochromatic arithmetic progression with m terms if

$$c \geq \left(1 + \frac{1}{m-2}\right)(m-2)^{\frac{1}{m-1}} N^{\frac{1}{m-1}}$$

which is to say that

$$W(c, m) > \frac{1}{\left(1 + \frac{1}{m-2}\right)^{m-2}} \cdot \frac{c^{m-1}}{m-1}$$

This bound is (slightly) better than the bound obtained with the LLL, i.e. Theorem 1.2 (see e.g. [12] and references therein) and, by virtue of the discussion above, the above bound is also better than that obtainable via the clique approximation of the CELL, i.e. Theorem 1.4.

3 Proof of Theorem 2.3

We start by noting that $|\mathcal{A}_l^\kappa(y) \cap \mathcal{S}_i| \leq d_l^\kappa$ (with d_l^κ the integer previously defined). Therefore the event A selected during procedure i_2) of the EC-algorithm described in the previous section is uniquely determined by a triple (l, κ, s) with $l \in E_{\mathcal{A}}$, $\kappa \in K_l$ and $s \in [d_l^\kappa]$. Observe also that at the end of procedure i_2) the variable y is always uncolored.

The EC-algorithm at each step i selects at random a number in the list $[k]$ and attributes this number to the smallest uncolored variable at step i . Let us suppose that these values are selected sequentially from the entries of a large vector $F_t \in [k]^t$ with $t \in \mathbb{N}$ sufficiently large. Hence $F_t = (f_1, \dots, f_t)$ with $f_j \in [k]$ for $j = 1, \dots, t$ and at step i of the algorithm the variable y to be evaluated receives the value $f_i \in [k]$ which is the entry i of F_t . We will show below that the vector F_t is uniquely determined by a “record” of the algorithm, namely a pair (R_t, C_t) where R_t is a string $r_1 \cdots r_i \cdots r_t$ and C_t is a function from \mathcal{V} to $[k] \cup \{0\}$.

Definition of the string $R_t = r_1 \cdots r_i \cdots r_t$. Suppose we are running the algorithm for t steps selecting the values of the variables sequentially from the entries of a fixed vector $F_t \in [k]^t$. For each step i of the algorithm define r_i as follows

- when at step i no bad event occur, put $r_i = 0$
- when at step i the bad event $A \in \mathcal{S}_i$, uniquely determined by the triple (l, κ, s) , is selected, then put $r_i = (l, \kappa, s)$ (we remind: $l \in E_{\mathcal{A}}$, $\kappa \in K_l$ and $s \in [d_l^\kappa]$).

Definition of the function C_t . Denote by $X_t \subset \mathcal{V}$ the set of uncolored variables after the step t (i.e. their values are still not assigned). If $x \in \mathcal{V} \setminus X_t$, then its value is assigned after t step and we call $C_t(x) \in [k]$ its assigned value. Let $C_t : \mathcal{V} \rightarrow [k] \cup \{0\}$ be the function that assigns to each variable $x \in \mathcal{V}$ either the value $C_t(x)$ if $x \in \mathcal{V} \setminus X_t$ or 0 if $x \in X_t$. So the function C_t tells us which variables are fixed (and at which values) and which are not fixed yet after t steps of the algorithm.

Observe that both R_t and C_t are uniquely determined by F_t via the set of instructions of the algorithm. We call \mathcal{M} the map $F_t \mapsto (R_t, C_t)$. The point is to prove that also the converse is true or, in other words, that \mathcal{M} is an injection. This means that the knowledge of the record (R_t, C_t) permits to reconstruct uniquely F_t and thus all the t steps made by the algorithm.

Lemma 3.1 *The set X_t is uniquely determined by the string R_t .*

Proof: We prove the statement by induction on t . At step $t = 1$ the first variable x (in the total order chosen) is colored and we have $R_1 = r_1$ with either $r_1 = 0$ when no bad event happens at step 1, or $r_1 = (1, 0, s)$ when a bad event A with $vbl(A) = \{x\}$ happens at step 1 (thus A has

size 1 and, being elementary, it such that $\kappa_x(A) = 0$ and $s \in [k]$ indicates the value of variable x which causes the occurrence of the event A . So $X_1 = \mathcal{V} \setminus \{x\}$ (when $r_1 = 0$) and $X_1 = \mathcal{V}$ (when $r_1 = (1, 0, s)$). Suppose now $t \geq 2$. By induction hypothesis X_{t-1} is uniquely determined by R_{t-1} . The smallest variable in the set X_{t-1} , call it y , is the variable that will be selected at step t . If $r_t = 0$, this means that the selection of the variable y has produced no bad events and so $X_t = X_{t-1} \setminus \{y\}$. If $r_t = (l, \kappa, s)$ then (l, κ, s) determines uniquely an event A among those of size l containing the variable y and we know that all variables in $vbl(A)$ have been uncolored except those in $G(A, y) \subset vbl(A)$ so that we have $X_t = X_{t-1} \setminus \{vbl(A) \setminus G(A, y)\}$. \square

Lemma 3.2 *For any $t \in \mathbb{N}$, the map \mathcal{M} that assigns to each vector $F \in [k]^t$ the pair (R_t, C_t) is injective.*

Proof: We prove by induction that the pair (R_t, C_t) uniquely determines F_t (in other words we prove that $|\mathcal{M}^{-1}(R, C_t)| = 1$). The claim is trivially true for $t = 1$. Indeed, given $R_1 = r_1$ with $r_1 = 0$, $C_1(x)$ is the value attributed to the first variable x . So the (one-dimensional) vector $F_1 = (f_1)$ has entry $f_1 = C_1(x)$. On the other hand, given $R_1 = r_1$ with $r_1 = (1, 0, s)$ we know that event $A = \{x = s\}$ has occurred and thus the (one-dimensional) vector $F_1 = (f_1)$ has entry $f_1 = s$. Suppose now the claim true for $t - 1$. Namely, the knowledge of the pair (R_{t-1}, C_{t-1}) implies the knowledge of the vector F_{t-1} . Therefore we have to show that, given the pair (R_t, C_t) we must be able to find the entry f_t of the vector F_t and the function C_{t-1} (so that (R_{t-1}, C_{t-1}) gives us all the remaining entries of the vector F_t).

We know through the Lemma 3.1 the set X_{t-1} and hence we know the smallest variable uncolored after $t - 1$ steps; call it y . We have to consider two cases: a) (R_t, C_t) is such that $r_t = 0$; b) (R_t, C_t) is such that $r_t = (l, \kappa, s)$. The case a) is easy. Indeed if $r_t = 0$ then the variable y will be colored at step t and so $f_t = C_t(y)$ while C_{t-1} is such that $C_{t-1}(x) = C_t(x)$ for all $x \in \mathcal{V} \setminus \{y\}$ and $C_{t-1}(y) = 0$. Let us now consider the case b) i.e. $r_t = (l, \kappa, s)$. Recall that the triple (l, κ, s) uniquely determines an event A . For this event A , occurring after the coloring of the variable y at the beginning of step t , we know the subset $G(A, y) \subset vbl(A)$ of the variables that will continue to stay colored after the conclusion of step t . Now we have C_t so we know the value of the κ variables belonging to $G(A, y)$. Since $G(A, y)$ is a seed of A , the knowledge of the variable in $G(A, y)$ (recall the definition given in (2.2)) allows us to deduce which was the value of all the other variables in $vbl(A) \setminus G(A, y)$. Call k_x^* the colors of the variable $x \in vbl(A)$ uniquely determined by the coloring of $G(A, y)$. Then $f_t = k_y^*$, while $C_{t-1}(x) = k_x^*$ when $x \in vbl(A) \setminus (G(A, y) \cup \{y\})$ and $C_{t-1}(x) = C_t(x)$ otherwise. \square

Let now denote by \mathcal{F}_t the subset of $[k]^t$ formed by all vectors F_t for which after t step of the algorithm there still are uncolored variables (i.e. the algorithm does not stop at step t). In other words let \mathcal{F}_t be the set of vectors F_t such that $C_t^{-1}(0) \neq \emptyset$. Clearly we have $|\mathcal{F}_t| \leq k^t$ and if we are able to prove that this inequality is strict, then this means that the set $[k]^t \setminus \mathcal{F}_t$ is non empty and for each vector $F_t \in [k]^t \setminus \mathcal{F}_t$, the algorithm stops.

Let $r \in [N]$. We further consider the set \mathcal{F}_t^r formed by all vectors F_t for which after t step of the algorithm there are exactly r variables not colored. In other words let \mathcal{F}_t^r be the set of vectors F_t such that $|C_t^{-1}(0)| = r$. Clearly \mathcal{F}_t is the disjoint union of the family $\{\mathcal{F}_t^r\}_{r \in [N]}$.

Let $(\mathcal{R}_t, \mathcal{C}_t)$ (resp. $(\mathcal{R}_t^r, \mathcal{C}_t^r)$) set of all record produced with vectors in \mathcal{F}_t (resp. \mathcal{F}_t^r), in other words $(\mathcal{R}_t, \mathcal{C}_t) = \mathcal{M}(\mathcal{F}_t)$ (resp. $(\mathcal{R}_t^r, \mathcal{C}_t^r) = \mathcal{M}(\mathcal{F}_t^r)$) is the image of \mathcal{F}_t through the map \mathcal{M} . We have, as a consequence of Lemma 3.2, the following proposition.

Proposition 3.3

$$|\mathcal{F}_t| \leq (k+1)^N \sum_{r=1}^N |\mathcal{R}_t^r| \quad (3.1)$$

Proof. Since \mathcal{F}_t is the disjoint union of the family $\{\mathcal{F}_t^r\}_{r \in [N]}$ we have that

$$|\mathcal{F}_t| = \sum_{r=1}^N |\mathcal{F}_t^r| \quad (3.2)$$

From Lemma 3.2 it follows that

$$|\mathcal{F}_t^r| \leq |\mathcal{R}_t^r| |\mathcal{C}_t^r| \leq |\mathcal{R}_t^r| |\mathcal{C}_t| \leq |\mathcal{R}_t^r| (k+1)^N \quad (3.3)$$

since \mathcal{C}_t is contained in the set of all functions from \mathcal{V} to $[k] \cup \{0\}$ which has cardinality $(k+1)^{|\mathcal{V}|} = (k+1)^N$. Putting (3.3) into (3.2) inequality (3.1) follows. \square

3.1 Upper bound for $|\mathcal{R}_t^r|$

A string $w_1 \dots w_n$ with $w_i \in \{0, 1\}$ is usually called a word on the alphabet $\{0, 1\}$. An initial segment of the word $w_1 \dots w_n$ is the a sub-string of $w_1 \dots w_n$ of the form $w_1 \dots w_i$ with $1 \leq i \leq n$. A *partial Dyck word* is word on the alphabet $\{0, 1\}$ such that in any initial segment of the word the number of 0's is greater or equal than the number of 1's. A *Dyck word* on the alphabet $\{0, 1\}$ is a partial Dyck word with equal number of 0's and 1's (hence a Dyck word has always an even number of letters). A partial Dyck word can be viewed as a path in \mathbb{Z}^2 starting at the origin made by steps either $(1, 1)$ or $(1, -1)$ in such a way that the path stays in the first quadrant (i.e. the path never goes below the x axis). A Dyck word (of size $2n$) is then a Dyck path which starts at the origin and ends at the point $(2n, 0)$ of the x axis. A descendant in a partial Dyck word is a sequence of 1's. For $S \subset \mathbb{N}$ and let us denote by $\mathcal{D}_{t,r,S}$ ($\mathcal{D}_{t,S}$) the set of partial Dyck words (Dyck words) with t 0's, $t-r$ 1's (t 1's) and with descendants having cardinality in S . Let

$$F = \{n \in \mathbb{N} : \text{exists } l \in E_{\mathcal{A}} \text{ and } \kappa \in K_l \text{ such that } n = l - \kappa\} \quad (3.4)$$

We now construct a (non injective) map $\mathfrak{M} : \mathcal{R}_t^r \rightarrow \mathcal{D}_{t,r,F}$.

Definition 3.4 Let $R_t = (r_1, \dots, r_t) \in \mathcal{R}_t^r$. We define the map \mathfrak{M} which associates to $R_t = (r_1, \dots, r_t)$ the word $R_t^* = r_1^* \dots r_t^*$ on the alphabet $\{0, 1\}$ as follows. If $r_i = 0$ then put $r_i^* = 0$. If $r_i = (l, \kappa, s)$ then put

$$r_i^* = 0 \overbrace{11 \dots 1}^{l-\kappa \text{ times}}$$

Lemma 3.5 For any $R_t \in \mathcal{R}_t^r$, the word R_t^* defined above is a partial Dyck word with t 0's and $t-r$ 1's and descendants in the set F defined in (3.4) above. In other words if $R_t \in \mathcal{R}_t^r$, then $R_t^* \in \mathcal{D}_{t,r,F}$.

Proof. Reading R_t^* from left to the right, every 0 correspond to a variable that has been colored and every 1 correspond to a variable that has been uncolored if some bad event occurs. So the number of 0's is t . Let x be the number of 1's. This number represents by construction the total number of variables that have been uncolored during the process of the algorithm. Thus $t-x = r$, i.e. $x = t-r$. The string is by construction a partial Dyck word since we cannot uncolor more variables than the number of colored variable. Finally again by construction we have that the descendants are of size $l-\kappa \in F$.

Lemma 3.6 *The pre-image of a string R_t^* by the map \mathfrak{M} has cardinality less or equal than*

$$\left[\max_{l, \kappa} \left\{ (d_l^\kappa)^{\frac{1}{l-\kappa}} \right\} \right]^{t-r} \prod_{j \in F} m_j^{D_j(R_t^*)}$$

where $D_j(R_t^*)$ is the number of descendants in R_t^* of size $l - \kappa = j$ and m_j is the multiplicity of $j = l - \kappa$, i.e.

$$m_j = \#\{(l, \kappa) : l - \kappa = j\}$$

Proof. We have to consider the information we lose in the map \mathfrak{M} . For each descent of size j in R_t^* , we have to specify the kind of event that generated this descent, for this, we have m_j possibilities. This accounts for the $\prod_{j \in F} m_j^{D_j(R_t^*)}$ factor. Once specified the kind of event of a given descent of size j , i.e., the pair (l, κ) , we have d_l^κ possible events that generated this descent. This accounts for a factor at most

$$\prod_{j \in F} \left[\max_{l, \kappa} d_l^\kappa \right]^{D_j(R_t^*)} = \prod_{j \in F} \left[\max_{l, \kappa} (d_l^\kappa)^{\frac{1}{l-\kappa}} \right]^{j D_j(R_t^*)} \leq \left[\max_{l, \kappa} \left\{ (d_l^\kappa)^{\frac{1}{l-\kappa}} \right\} \right]^{t-r},$$

since $\sum_{j \in F} j D_j(R_t^*) = t - r$. \square

Remark. Note that

$$\phi_E(x) = 1 + \sum_{l \in E} \sum_{\kappa \in K_l} x^{l-\kappa} = 1 + \sum_{j \in F} m_j x^j \quad (3.5)$$

We now have the following lemmas (Lemma 6 and Lemma 7 in [10]).

Lemma 3.7 *The number of partial Dyck words in $\mathcal{D}_{t,r,F}$ is less than the number of Dyck words in $\mathcal{D}_{t+r(s-1),F}$ where $s = \min\{F \setminus \{1\}\}$.*

Lemma 3.8 *There is a bijection between Dyck words in $\mathcal{D}_{t+r(s-1),F}$ and plane rooted trees on $t + r(s-1) + 1$ vertices and with degree (number of children in the vertices) in $F \cup \{0\}$.*

Using these last three lemmas we arrive at the following proposition.

Proposition 3.9 *Let \mathcal{T}_n denotes the set of all plane rooted trees with n vertices, then we have*

$$|\mathcal{R}_t^r| \leq \left[\max_{l, \kappa} \left\{ (d_l^\kappa)^{\frac{1}{l-\kappa}} \right\} \right]^{t-r} \sum_{T \in \mathcal{T}_{t+r(s-1)+1}} \prod_{j \geq 0} \varphi_j^{D_j(T)} \quad (3.6)$$

where $D_j(T)$ denotes the number of vertices with j children in the tree T and

$$\varphi_j = \begin{cases} 1 & \text{if } j = 0 \\ m_j & \text{if } j \in F \\ 0 & \text{otherwise} \end{cases} \quad (3.7)$$

Proof. The proof follows from the three lemmas above and the observation that

$$\sum_{T \in \mathcal{T}_{t+r(s-1)+1}} \omega(T) = \sum_{R^* \in \mathcal{D}_{t+r(s-1),F}} \prod_{j \in F} m_j^{D_j(R_t^*)}$$

We finally use the following theorem (Theorem 5 in [7]).

Theorem 3.10 *Let $\varphi(x) = \sum_{j \geq 0} \varphi_j x^j$, let R be the convergence radius of $\varphi(x)$ and let τ be the first root in $[0, R)$ of the equation $x\varphi'(x) - \varphi(x) = 0$. Set $d = \gcd\{j > 0 : \varphi_j > 0\}$. Then*

$$\sum_{T \in \mathcal{T}_n} \prod_{j \geq 0} \varphi_j^{D_j(T)} = d \left[\frac{\varphi(\tau)}{2\pi\varphi''(\tau)} \right] \frac{[\varphi'(\tau)]^n}{n^{3/2}} (1 + \mathcal{O}(n^{-1})) \leq C_\varphi \frac{[\varphi'(\tau)]^n}{n^{3/2}} \quad (3.8)$$

with C_φ constant.

Inserting now inequality (3.8) into (3.6) and recalling (3.5) we get

$$|\mathcal{R}_t^r| \leq \left[\max_{l, \kappa} \left\{ (d_l^\kappa)^{\frac{1}{l-\kappa}} \right\} \right]^{t-r} C_\varphi \frac{[\phi'_E(\tau)]^{t+r(s-1)+1}}{(t+r(s-1)+1)^{3/2}} \quad (3.9)$$

where C_φ is the constant appearing in Theorem 3.10 and φ_j is defined as in (3.7).

3.2 Conclusion of the proof of Theorem 2.3

We are now in the position to end the proof of Theorem 2.3. Indeed, inserting (3.9) into (3.1) we get

$$\begin{aligned} |\mathcal{F}_t| &\leq C_\varphi (k+1)^N \sum_{r=1}^N \left[\max_{l, \kappa} \left\{ (d_l^\kappa)^{\frac{1}{l-\kappa}} \right\} \right]^{t-r} \frac{[\phi'_E(\tau)]^{t+r(s-1)+1}}{(t+r(s-1)+1)^{3/2}} \leq \\ &\leq C(E, N, k) \left[\max_{l, \kappa} \left\{ (d_l^\kappa)^{\frac{1}{l-\kappa}} \right\} \cdot \phi'_E(\tau) \right]^t \frac{1}{t^{3/2}} \end{aligned}$$

with

$$C(E, N, k) \leq C_\varphi (k+1)^N [\phi'_E(\tau)]^{N(s-1)+1}$$

Hence if

$$k \geq \left[\max_{l, \kappa} \left\{ (d_l^\kappa)^{\frac{1}{l-\kappa}} \right\} \cdot \phi'_E(\tau) \right] \quad (3.10)$$

we have that for t sufficiently large

$$|\mathcal{F}_t| < k^t$$

i.e. the algorithm stops. Note that the condition (3.10) which ensures the algorithm to stop is exactly (2.6) with greater or equal replacing strictly greater. Now note that $|\mathcal{F}_t|/k^t$ is the probability that after t steps the algorithm is still running. Let us suppose that inequality (3.10) holds strictly. I.e., let us suppose that there exists $\varepsilon > 0$ such that $(\phi'_E(\tau)/k) \cdot \max_{l, \kappa} (d_l^\kappa)^{1/(l-\kappa)} \leq (1-\varepsilon)$. Then the probability $P(t)$ that the algorithm is still running after t steps can be bounded above as follows

$$P(t) \leq C_\varphi (k+1)^N [\phi'_E(\tau)]^{N(s-1)+1} (1-\varepsilon)^t$$

Let t_0 be the solution of the equation

$$C_\varphi (k+1)^N [\phi'_E(\tau)]^{N(s-1)+1} (1-\varepsilon)^t = 1$$

Remark that t_0 is linear in N and observe that $P(t_0 + t) \leq (1-\varepsilon)^t = e^{-t|\ln(1-\varepsilon)|}$. Therefore, if inequality (3.10) holds strictly, the expected running time T of the EC-algorithm is bounded by

$$T \leq t_0 + \frac{1}{|\ln(1-\varepsilon)|}$$

This concludes the proof of Theorem 2.3.

So if the condition (3.10) is replaced by the inequality

$$\left[\max_{l, \kappa} \left\{ (d_l^\kappa)^{\frac{1}{l-\kappa}} \right\} \cdot \phi'_E(\tau) \right] \leq (1 - \varepsilon)k \quad (3.11)$$

for some $\varepsilon > 0$, then the expected running time of the algorithm is linear in N .

4 An application: coloring graphs frugally

Given a graph $G = (V, E)$, a coloring of the vertices of G is *proper* if no two adjacent vertices receive the same color and a proper vertex coloring of G is β -*frugal* if any vertex has at most β members of any color class in its neighborhood. The minimum number of colors required such that a graph G has at least one β -frugal proper vertex coloring is called the β -*frugal chromatic number* of G and will be denoted by $\chi_\beta(G)$.

Theorem 4.1 *Let $G = (V, E)$ be a graph with maximum degree Δ . Then, for any $\beta \in [\Delta - 1]$ we have*

$$\chi_\beta(G) \leq q(\beta)\Delta \max \left\{ 1, \left(\frac{\Delta}{\beta!} \right)^{\frac{1}{\beta}} \right\} \quad (4.1)$$

with

$$q(\beta) = \inf_{x>0} \left[\frac{1+x+x^\beta}{x} \right]$$

Proof. Let $G = (V, E)$ be a graph with vertex set V , edge set E and with maximum degree Δ . Suppose that we color the vertices of G by choosing at random independently and uniformly from a set C of k colors and choose a bijection $f : [k] \rightarrow C$.

We are thus in the i.i.d. variable setting in which the set i.i.d. random variables is $\mathcal{V} = \{x_v\}_{v \in V}$ are indexed by the set of the vertices V of G and each variable x_v takes values in $[k]$ such that, for $j \in [k]$, $x_v = j$ means that the vertex v is colored with the color $f(j) \in C$.

Given $v \in V$, we denote by H_β^v the set of all subsets of the neighborhood $\Gamma^*(v)$ of v with exactly $\beta + 1$ vertices. In other words

$$H_\beta^v = \{Y \subset \Gamma^*(v) : |Y| = \beta + 1\}$$

So an element $h_\beta^v \in H_\beta^v$ is a set of $\beta + 1$ vertices of $\Gamma_G^*(v)$ (i.e. $h_\beta^v = \{v_1, v_2, \dots, v_{\beta+1}\} \subset \Gamma_G^*(v)$).

- Given $e = \{u, v\} \in E$, let A_e be the event that u and v receive the same color (i.e. e is monochromatic).

- Given $v \in V$ and $h_\beta^v \in H_\beta^v$, we let $A_{h_\beta^v}$ to be the event all the vertices of the set h_β^v receive the same color (i.e. h_β^v is monochromatic).

We thus have a family of (bad) events $\mathcal{A} = \{A_e\}_{e \in E} \cup \{A_{h_\beta^v}\}_{v \in V, h_\beta^v \in H_\beta^v}$.

Theorem 2.3 tells us that when condition (2.6) is satisfied, then it is possible to find a coloring (i.e. an outcome in the sample space generated by the variables \mathcal{V}) avoiding all bad events of the family \mathcal{A} . I.e. it is possible to find a β -frugal coloring of the vertex set V of G . Therefore, if k is greater or equal to the r.h.s of (2.6) then G admits a β -frugal coloring and hence $\chi^\beta(G) \leq k$.

To check thus condition (2.6) let us observe that:

- If $e = \{v, w\} \in E$ is an edge of G , then $vbl(A_e) = \{x_v, x_w\}$ and thus $|vbl(A_e)| = 2$;

- If $h_\beta^v = \{v_1, \dots, v_{\beta+1}\}$ is a subset of $\Gamma^*(v)$ for some $v \in V$, then $vbl(A_{h_\beta^v}) = \{x_{v_1}, \dots, x_{v_{\beta+1}}\}$ and thus $|vbl(A_{h_\beta^v})| = \beta + 1$.

Therefore

$$E_{\mathcal{A}} = \{2, \beta + 1\} \quad \text{and} \quad \mathcal{A}_2 = \{A_e\}_{e \in \mathcal{E}} \quad \text{and} \quad \mathcal{A}_{\beta+1} = \{A_{h_\beta^v}\}_{v \in V, h_\beta^v \in H_\beta^v}$$

Moreover,

- for any $v \in V$ and for any $A \in \mathcal{A}_2(x_v)$ we have that $\kappa_{x_v}(A) = 1$ and hence $K_2 = \{1\}$;
- for any $v \in V$ and for any $A \in \mathcal{A}_{\beta+1}(x_v)$ we have that $\kappa_{x_v}(A) = 1$ and hence $K_{\beta+1} = \{1\}$;

Therefore the function $\phi_{E_{\mathcal{A}}}(x)$ defined in (2.4) is in the present case

$$\phi_{E_{\mathcal{A}}}(x) = 1 + x + x^\beta$$

Let us pose

$$q(\beta) = \inf_{x>0} \left[\frac{1 + x + x^\beta}{x} \right] \quad (4.2)$$

Let us now estimate the numbers d_2^1 and $d_{\beta+1}^1$. Since d_2^1 represents the maximum number of edges of G containing a fixed vertex, we have

$$d_2^1 \leq \Delta \quad (4.3)$$

Concerning the estimate of $d_{\beta+1}^1$, observe that a vertex $w \in V$ has at most Δ neighbors $\{v_1, \dots, v_\Delta\}$ and therefore w belongs to at most Δ distinct neighborhoods $\Gamma^*(v_1), \dots, \Gamma^*(v_\Delta)$ and in each $\Gamma^*(v_i)$ of these neighborhoods w can be contained in at most $\binom{\Delta}{\beta}$ sets of type $h_\beta^{v_i} = \{v_1^i, v_2^i, \dots, v_{\beta+1}^i\}$. Therefore

$$d_{\beta+1}^1 \leq \Delta \binom{\Delta}{\beta} \leq \Delta^{1+\beta} / \beta! \quad (4.4)$$

We can now check condition (2.6). Namely, there is a β -frugal coloring of G as soon as

$$k \geq q(\beta) \max \left\{ (d_2^1)^{\frac{1}{2-1}}, (d_{\beta+1}^1)^{\frac{1}{\beta+1-1}} \right\} \quad (4.5)$$

Due to inequalities (4.3) and (4.4), condition (4.5) is surely satisfied if

$$k \geq q(\beta) \max \left\{ \Delta, \frac{\Delta^{1+\frac{1}{\beta}}}{(\beta!)^{1/\beta}} \right\} \quad (4.6)$$

and hence

$$\chi_\beta(G) \leq q(\beta) \max \left\{ \Delta, \frac{\Delta^{1+\frac{1}{\beta}}}{(\beta!)^{1/\beta}} \right\} \quad (4.7)$$

which is the bound (4.1). \square

The bound (4.1) represents an improvement on the previous best bound for the β -frugal chromatic number of a graph with maximum degree Δ given in [18] (see Sec. 2.6, formula (2.18) in [18] and compare $q(\beta)$ defined in (4.2) above with the functions $k_1(\beta)$ and $k_2(\beta)$ there defined:

one can check that $q(\beta) \leq \min\{k_1(\beta), k_2(\beta)\}$. In particular, for the case of 2-frugal coloring (a.k.a. linear coloring) we get

$$\chi_2(G) \leq \frac{3}{\sqrt{2}} \Delta^{\frac{3}{2}}$$

In [24] it is proved that $\Delta^{\frac{3}{2}}$ is the correct order since by an explicit example the author proves that $\chi_2(G) \geq \Delta^{3/2}/(6\sqrt{3})$ while his upper bound (using the LLL) is $\chi_2(G) \leq 10\Delta^{\frac{3}{2}}$ and the improved upper bound given in [18] is $\chi_2(G) \leq \frac{4.9127}{\sqrt{2}} \Delta^{3/2}$.

Acknowledgments

A.P. and R.S. have been partially supported by the Brazilian agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG - Programa de Pesquisador Mineiro).

References

- [1] Alon, N.: *A parallel algorithmic version of the local lemma*. Random Structures and Algorithms, **2**, n. 4, 367-378 (1991).
- [2] Alves, R. G.; Procacci, A.: *Witness trees in the Moser-Tardos algorithmic Lovász Local Lemma and Penrose trees in the hard-core lattice gas*, Journal of Statistical Physics, **156**, p. 877-895 (2014)
- [3] Beck, J.: *An Algorithmic Approach to the Lovsz Local Lemma*, Random Structures and Algorithms, **2**, n. 4, 343-365 (1991).
- [4] Bissacot, R.; Fernández, R.; Procacci A.; Scoppola, B.: *An Improvement of the Lovász Local Lemma via Cluster Expansion*, Combinatorics Probability and Computing, **20**, n. 5, 709-719 (2011)
- [5] Böttcher, J.; Kohayakawa, Y.; Procacci, A.: *Properly coloured copies and rainbow copies of large graphs with small maximum degree*, Random Structures and Algorithms, **40**, n. 4, 425-436 (2012).
- [6] Dujmovic, V.; Joret, G.; Kozik, J.; Wood, D. R.: *Nonrepetitive colouring via entropy compression*, to appear in Combinatorica, DOI: 10.1007/s00493-015-3070-6.
- [7] Drmota, M.: *Combinatorics and asymptotics on trees*, Cubo J. **6** (2) (2004).
- [8] Dobrushin, R. L.: *Perturbation methods of the theory of Gibbsian fields*, in P. Bernard (editor), Lectures on Probability Theory and Statistics, Lecture Notes in Mathematics Volume 1648, 1996, pp 1-66 Springer-Verlag, Berlin, (1996).
- [9] Erdős, P. and Lovász, L.: *Problems and results on 3-chromatic hypergraphs and some related questions, in Infinite and finite sets*. Vol. II, Colloq. Math. Soc. Janos Bolyai, Vol. 10, pp. 609-627. North-Holland, Amsterdam, (1975).
- [10] Esperet, L.; Parreau, A.: *Acyclic edge-coloring using entropy compression*, European Journal of Combinatorics, **34**, In. 6, 1019- 1027 (2013).

- [11] Fernández, R.; Procacci A.: *Cluster expansion for abstract polymer models. New bounds from an old approach*, Communications in Mathematical Physics. **274**, n.1, 123-140 (2007).
- [12] Gasarch W.; Haeupler B.: *Lower Bounds on van der Waerden Numbers: Randomized and Deterministic-Constructive*, The Electronic J. Combinatorics, **18** (2011), #P64.
- [13] Kolipaka, K. B. R.; Szegedy, M., *Moser and Tardos meet Lovász*, Proceedings of the 43rd annual ACM symposium on Theory of computing Pages 235-244, ACM New York, NY, USA (2011).
- [14] McDiarmid. C.: *Hypergraph coloring and the Lovász Local Lemma*, Discrete Math., **167/168**, 481-486 (1997).
- [15] Moser, R. A.: *A constructive proof of the Lovász Local Lemma*, in Proceedings of the 41st Annual ACM Symposium on the Theory of Computing (STOC). ACM, New York (2009).
- [16] Moser, R. A.: *Exact algorithms for constraint satisfaction problems*, Diss. ETH Zurich, Nr. 20668, Logos Verlag, Berlin (2013).
- [17] Moser, R. ; Tardos, G.: *A constructive proof of the general Lovász Local Lemma*, J. ACM **57** Article 11, 15 pages (2010).
- [18] Ndreca, S.; Procacci, A.; Scoppola, B.: *Improved bounds on coloring of graphs*, European Journal of Combinatorics, **33**, n 4, p. 592-609 (2012).
- [19] Pegden, W.: *An extension of the Moser-Tardos algorithmic local lemma*, SIAM J. Discrete Math. **28**, 911-917 (2013).
- [20] Procacci, A.; Sanchis, R.: *Perfect and separating hash families: new bounds via the algorithmic cluster expansion local lemma*, Annales de l'Institut Henry Poincaré D combinatorics, physics and their interactions, to appear (2016).
- [21] Przybylo, J.: *On the Facial Thue Choice Index via Entropy Compression*, Journal of Graph Theory, **77**, Issue 3, 180-189, (2014).
- [22] Scott, A.; Sokal, A. D.: *The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma*, J. Stat. Phys. **118**, no. 5-6, 1151–1261, (2005).
- [23] Tao, T.: *Moser's entropy compression argument*, Terence Tao Blog post, (2009).
- [24] Yuster, R.: *Linear coloring of graphs*, Discr. Math. **185**, Issues 13, Pages 293-297 (1998).